L Number	Hits	Search Text	DB	Time stamp
1	125	705/33, 34.ccls. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:05
3	0	(705/33, 34.ccls. and (duplicate near2 (record$1 or account$1))) and ((compar$5 near3 (index$2 or account$1)) with replac$3)	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:03
2	35	(705/33, 34.ccls. and (duplicate near2 (record$1 or account$1))) and compar$5 and replac$3	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:04
4	0	705/$.ab,clm,ab. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:05
5	0	705/ab,clm,ab.ccls. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:06
6	0	705/ab,clm,ab. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:13
7	0	705/ab,clm,ab.ccls. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:13
8	0	705/ab,clm,ti.ccls. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:13
9	0	705/ab,ti.ccls. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:13
10	146	705/$.ccls.. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:14

11	146	705/$.ccls. and (duplicate near2 (record$1 or account$1))	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:14
12	4	705/$.ccls. and ((duplicate near2 (record$1 or account$1)) same replac$3)	USPAT; US-PGPUB; EPO; JPO; DERWENT; IBM_TDB	2004/09/20 09:14

STIC Search Report
EIC 2100

USPTO

STIC Database Tracking Number: 132943

TO: Baoquoc To **Location:** **Art Unit : 2172** **Monday, September 20, 2004** **Case Serial Number: 09/832572**	**From: David Holloway** **Location: EIC 2100** **PK2-4B30** **Phone: 308-7794** **david.holloway@uspto.gov**

Search Notes

Dear Examiner To,

Attached please find your search results for above-referenced case.
Please contact me if you have any questions or would like a re-focused search.

David

sira
Search and Information Resources Administration

STIC EIC 2100
Search Request Form

Today's Date: 09/20/04

What date would you like to use to limit the search?
Priority Date: 04/01/2001 Other:

Name TO, BAO QUOC
AU 2172 Examiner # Boss 78889
Room # 4A42 Phone 305 9469
Serial # 09/832,572

Format for Search Results (Circle One):
PAPER DISK EMAIL

Where have you searched so far?
USP DWPI EPO JPO ACM IBM TDB
IEEE INSPEC SPI Other _____

Is this a "Fast & Focused" Search Request? (Circle One) YES NO
A "Fast & Focused" Search is completed in 2-3 hours (maximum). The search must be on a very specific topic and meet certain criteria. The criteria are posted in EIC2100 and on the EIC2100 NPL Web Page at http://ptoweb/patents/stic/stic-tc2100.htm.

What is the topic, novelty, motivation, utility, or other specific details defining the desired focus of this search? Please include the concepts, synonyms, keywords, acronyms, definitions, strategies, and anything else that helps to describe the topic. Please attach a copy of the abstract, background, brief summary, pertinent claims and any citations of relevant art you have found.

- identify duplicate and eliminating duplicate
- if the field is null enter date
- Claims 1, 6, 7

STIC Searcher David Holloway Phone 308-7794
Date picked up 9-20-04 Date Completed 9-20-04

Search and Information Resources Administration

```
Set     Items   Description
S1     957749   DATABASE? OR DATABANK? OR DB OR DBMS OR RDB OR OODB OR DBS
           OR DATA()(BASE? OR BANK?)
S2    2265367   RECORD? OR DOCUMENT? OR INVOICE? OR BILL OR BILLS OR RECEI-
           PT?
S3    2877079   DUPLICATE OR IDENTICAL OR DUPLICATES OR SAME OR EQUIVALENT
S4    2318766   REMOVE? OR REPLAC? OR SUBSTITUT? OR SWITCH? OR SWAP?
S5    5820849   DATE OR DATES OR DATESTAMP?  OR INDICIA? OR TIMESTAMP? OR -
           TIME OR CALENDAR?
S6     319942   BLANK? OR NULL OR EMPTY OR VOID?
S7    5425155   MATCH? OR COMPARE? OR EQUAL OR EQUIVALENT?
S8      18621   S2(2N)(S3 OR S7)
S9       1674   S1 AND S8
S10        90   S4 AND S9
S11       942   S1 AND S5 AND S6
S12     13473   S5(2N)(INSERT? OR FILL OR FEED OR POPULAT?)
S13        18   S11 AND S12
S14        60   S12(5N)S6
S15   1584109   BLEND OR BLENDS OR BLENDING OR MERGE? OR MERGING OR COMBINE
           OR COMBINES OR COMBINING OR APPEND? OR MIX OR UPDAT? OR UP()-
           (DATE OR DATING OR DATES)
S16        93   S11 AND S15
S17         1   S14 AND S15
S18       165   S10 OR S13 OR S14 OR S17
S19       110   RD (unique items)
S20        88   S19 NOT PY>2001
S21        86   S20 NOT PD>20010412
S22        10   S21 AND S15
S23       108   S13 OR S10
S24        77   RD (unique items)
S25        63   S24 NOT PY>2001
S26        61   S25 NOT PD>20010412
S27        52   S26 NOT S22
File   8:Ei Compendex(R) 1970-2004/Sep W2
         (c) 2004 Elsevier Eng.  Info. Inc.
File  35:Dissertation Abs Online 1861-2004/Aug
         (c) 2004 ProQuest Info&Learning
File 202:Info. Sci. & Tech. Abs. 1966-2004/Sep 09
         (c) 2004 EBSCO Publishing
File  65:Inside Conferences 1993-2004/Sep W3
         (c) 2004 BLDSC all rts. reserv.
File   2:INSPEC 1969-2004/Sep W2
         (c) 2004 Institution of Electrical Engineers
File  94:JICST-EPlus 1985-2004/Aug W4
         (c)2004 Japan Science and Tech Corp(JST)
File 111:TGG Natl.Newspaper Index(SM) 1979-2004/Sep 20
         (c) 2004 The Gale Group
File 233:Internet & Personal Comp. Abs. 1981-2003/Sep
         (c) 2003 EBSCO Pub.
File   6:NTIS 1964-2004/Sep W2
         (c) 2004 NTIS, Intl Cpyrght All Rights Res
File 144:Pascal 1973-2004/Sep W2
         (c) 2004 INIST/CNRS
File  34:SciSearch(R) Cited Ref Sci 1990-2004/Sep W2
         (c) 2004 Inst for Sci Info
File  99:Wilson Appl. Sci & Tech Abs 1983-2004/Aug
         (c) 2004 The HW Wilson Co.
File  95:TEME-Technology & Management 1989-2004/Jun W1
         (c) 2004 FIZ TECHNIK
```

22/5/2 **(Item 1 from file: 202)**

2800798
Bibliography formatting software: an update .
Author(s): Stigleman, S
Database vol. 16, no. 1, pages 24-37
Publication Date: Feb 1993
ISSN: 0162-4105
Language: English
Document Type: Journal Article
Record Type: Abstract
Journal Announcement: 2800

This article provides an **update** on bibliography formatting software on the market, noting the currently available 52 programs. The author also reports on what has changed in the 43 programs that were available in 1992. Some features appearing in increased numbers of programs are studied, including detection of **duplicate records** , moving or copying blocks, searching and **replacing** call number sorting, glossaries for frequently entered text, pop-up field contents, and dictionaries of journal abbreviations. Importing records obtained from online or CD-ROM **databases** is seen as a major driving force in the development of bibliographic software.

Descriptors: Bibliographic systems; Bibliographies; Citations; Computer programs
Classification Codes and Description: 5.06 (Software and Programming); 6.02 (Bibliographic Search Services, **Databases**)
Main Heading: Information Processing and Control; Information Systems and Applications

22/5/5 **(Item 1 from file: 233)**
DIALOG(R)File 233:Internet & Personal Comp. Abs.

00604110 00IT06-005
Ovid launches new version of search-and-retrieval software
Information Today , June 1, 2000 , v17 n6 p74-75, 2 Page(s)
ISSN: 8755-6286
Company Name: Ovid Technologies
URL: http://www.ovid.com
Product Name: Ovid 4.1.0
Languages: English
Document Type: Product Announcement
Geographic Location: United States

Announces that Ovid Technologies, Inc. of New York, NY (800) will release version 4.1.0 of its search-and-retrieval software. Says that the Multifile and Deduping features allow users to **combine** multiple Ovid **databases** , search them simultaneously, and automatically **remove duplicate records** . Notes that enhancements to its WebLinks technology allow sites to define and customize links from Ovid records to non-Ovid, Web-based resources. Details the capabilities of the Multifile and Deduping features, and the changes to WebLinks customization options. Mentions that Ovid is developing an OpenLinks service for creating links from its bibliographic **database** records to specific journal articles on publisher Web sites. Highlights nine other new features, from an option to delete searches to obtaining a history of the full-text documents accessed. (amg)

Descriptors: Online Information; Software Tools; **Database** ; Product Development
Identifiers: Ovid 4.1.0; Ovid Technologies

27/5/10 **(Item 1 from file: 202)**
DIALOG(R)File 202:Info. Sci. & Tech. Abs.

1701204
Automated matching and amalgamation of marc records in the dobis database
.
Author(s): Mouland, P; Webber, R
Corporate Source: Library Systems Centre, National Library of Canada
Canadian Journal of Information Science vol. 6, pages 57-65
Publication Date: June 1981
ISSN: 1195-096X
Language: English
Document Type: Journal Article
Record Type: Abstract
Journal Announcement: 1700

This paper outlines the techniques used at the national library of canada for records loaded off-line to achieve the objectives that there be a single bibliographic record for each unique work and that any changes to these records improve overall quality of the **database** . The authors note the criteria used to identify **duplicate records** and describe the process of deciding whether to **replace** a record completely or to perform record amalgamation. The amalgamation process is outlined, and features of the system which allow the on-line cataloguer to protect data elements from off-line modification are explained

Classification Codes and Description: 6.02 (Bibliographic Search Services, **Databases**); 2.01 (Definitions, Theoretical Considerations
Main Heading: Information Systems and Applications; Research Methods

27/5/11 (Item 1 from file: 2)
DIALOG(R)File 2:INSPEC

6576736 INSPEC Abstract Number: C2000-06-6160Z-008
Title: Cleansing data for mining and warehousing
Author(s): Mong Li Lee; Hongjun Lu; Tok Wang Ling; Yee Teng Ko
Author Affiliation: Sch. of Comput., Nat. Univ. of Singapore, Singapore
Conference Title: Database and Expert Systems Applications. 10th International Conference, DEXA'99 (Lecture Notes in Computer Science Vol.1677) p.751-60
Editor(s): Bench-Capon, T.; Soda, G.; Tjoa, A.M.
Publisher: Springer-Verlag, Berlin, Germany
Publication Date: 1999 Country of Publication: Germany xviii+1105 pp.
ISBN: 3 540 66448 3 Material Identity Number: XX-1999-02591
Conference Title: Proceedings of DEXA'99: 10th International Conference and Workshop on Database and Expert Systems Applications
Conference Date: 30 Aug.-3 Sept. 1999 Conference Location: Florence, Italy
Language: English Document Type: Conference Paper (PA)
Treatment: Theoretical (T)
Abstract: Given the rapid growth of data, it is important to extract, mine and discover useful information from **databases** and data warehouses. The process of data cleansing is crucial because of the "garbage in, garbage out" principle. "Dirty" data files are prevalent because of incorrect or missing data values, inconsistent value naming conventions, and incomplete information. Hence, we may have multiple records referring to the same real-world entity. We examine the problem of detecting and removing duplicating records. We present several efficient techniques to pre-process the records before sorting them so that potentially **matching records** will be brought to a close neighbourhood. Based on these techniques, we implement a data cleansing system which can detect and **remove** more **duplicate records** than existing methods. (7 Refs)
Subfile: C
Descriptors: data integrity; data mining; data warehouses; **database** theory
Identifiers: data mining; information discovery; data warehouses; data cleansing; data inconsistency; value naming conventions; missing data values; incomplete information; record pre-processing; **duplicate record** removal
Class Codes: C6160Z (Other DBMS); C4250 (Database theory); C6130 (Data handling techniques)

27/5/16 **(Item 6 from file: 2)**
DIALOG(R)File 2:INSPEC

5592775 INSPEC Abstract Number: C9707-6130D-004
Title: Duplicate document **detection**
Author(s): Spitz, A.L.
Author Affiliation: Daimler Benz Res. & Technol. Center, Palo Alto, CA, USA
Journal: Proceedings of the SPIE - The International Society for Optical Engineering Conference Title: Proc. SPIE - Int. Soc. Opt. Eng. (USA) vol.3027 p.88-94
Publisher: SPIE-Int. Soc. Opt. Eng,
Publication Date: 1997 Country of Publication: USA
CODEN: PSISDG ISSN: 0277-786X
SICI: 0277-786X(1997)3027L.88:DDD;1-U
Material Identity Number: C574-97101
U.S. Copyright Clearance Center Code: 0 8194 2438 2/97/$10.00
Conference Title: Document Recognition IV
Conference Sponsor: SPIE; Soc. Imaging Sci. & Technol
Conference Date: 12-13 Feb. 1997 Conference Location: San Jose, CA, USA
Language: English Document Type: Conference Paper (PA); Journal Paper (JP)
Treatment: Practical (P)
Abstract: In document image filing applications it is important to be able to recognize whether a particular document has already been entered into the system either as an individual document or as an inclusion in another document. Document images could be matched on the basis of layout or contents. However, matching of layout may not be effective when style is strictly controlled. We develop a document "handle" which is stored along with the document image. The handle is simply a character shape coded representation of the image after the figures and tables have been **removed**. Character shape coding is a method of identifying individual character images as members of one of a small number of classes. This process is computationally inexpensive and tolerant of differing generations of photocopying, skew and scanner characteristics. When a new document is entered into the system, its handle is computed and compared against all of the extant handles using a normalized Levenshtein metric. We demonstrate the ability to detect **duplicate documents** comprising single and multiple pages. (6 Refs)
Subfile: C
Descriptors: document image processing; image coding; image matching; image representation; optical character recognition; visual **databases**
Identifiers: **duplicate document** detection; document image filing applications; **document** image **matching** ; document layout; document handle ; character shape coded representation; tables; computationally inexpensive ; photocopying; skew; scanner; normalized Levenshtein metric
Class Codes: C6130D (Document processing techniques); C6160S (Spatial and pictorial databases); C5260B (Computer vision and image processing techniques); C1250 (Pattern recognition)

27/5/19 (Item 9 from file: 2)
DIALOG(R)File 2:INSPEC

4777738 INSPEC Abstract Number: C9411-6160K-007
Title: Techniques for indexing large numbers of constraints and rules in a database system
Author(s): Kumar, A.
Author Affiliation: Graduate Sch. of Manage., Cornell Univ., Ithaca, NY, USA
p.65-71
Editor(s): Tjoa, A.M.; Ramos, I.
Publisher: Springer-Verlag, Wien, Austria
Publication Date: 1992 Country of Publication: Austria xii+546 pp.
ISBN: 3 211 82400 6
Conference Title: Proceedings of DEXA '92. International Conference on Database and Expert Systems Applications
Conference Date: 2-4 Sept. 1992 Conference Location: Valencia, Spain
Language: English Document Type: Conference Paper (PA)
Treatment: Practical (P)
Abstract: Addresses the problem of indexing a large number of rules and constraints in a **database** system. The objective of such indexing is to be able to quickly identify the relevant constraints and rules, rather than search sequentially every **time** **insertions** , deletions and modifications are made to the **database** . The constraints are represented as SQL queries which must return **null** answers. Each constraint is parsed and stored in one or more indexes. Algorithms for index maintenance and constraint retrieval are given. (16 Refs)
Subfile: C
Descriptors: constraint handling; deductive **databases** ; indexing; relational **databases** ; SQL
Identifiers: indexing techniques; relevant constraints; rules; **database** system; insertions; deletions; modifications; SQL queries; **null** answers; parsing; index maintenance algorithms; constraint retrieval
Class Codes: C6160K (Deductive databases); C6160D (Relational DBMS)

```
Set      Items   Description
S1      163084   DATABASE? OR DATABANK? OR DB OR DBMS OR RDB OR OODB OR DBS
             OR DATA()(BASE? OR BANK?)
S2     1144561   RECORD? OR DOCUMENT? OR INVOICE? OR BILL OR BILLS OR RECEI-
             PT?
S3     1648187   DUPLICATE OR IDENTICAL OR DUPLICATES OR SAME OR EQUIVALENT
S4     2716345   REMOVE? OR REPLAC? OR SUBSTITUT? OR SWITCH? OR SWAP?
S5     2724453   DATE OR DATES OR DATESTAMP?  OR INDICIA? OR TIMESTAMP? OR -
            TIME OR CALENDAR?
S6      239544   BLANK? OR NULL OR EMPTY OR VOID?
S7     1302872   MATCH? OR COMPARE? OR EQUAL OR EQUIVALENT?
S8       15609   S2(2N)(S3 OR S7)
S9         992   S1 AND S8
S10         64   S4 AND S9
S11        296   S1 AND S5 AND S6
S12      12638   S5(2N)(INSERT? OR FILL OR FEED OR POPULAT?)
S13          1   S11 AND S12
S14         80   S12(5N)S6
S15         46   (S10 OR S13 OR S14) AND IC=G06F?
S16         36   S15 NOT AD>20010412
S17         36   IDPAT (sorted in duplicate/non-duplicate order)
S18         35   IDPAT (primary/non-duplicate records only)
File 347:JAPIO Nov 1976-2004/May(Updated 040903)
         (c) 2004 JPO & JAPIO
File 350:Derwent WPIX 1963-2004/UD,UM &UP=200459
         (c) 2004  Thomson Derwent
```

18/5/1 **(Item 1 from file: 350)**
DIALOG(R)File 350:Derwent WPIX

016135112 **Image available**
WPI Acc No: 2004-292988/200427
Related WPI Acc No: 2003-707698
XRPX Acc No: N04-232577

Multi-copies synchronizing method of database **, involves identifying edited and modified records by comparing records of focus copy against** records **having** same **identification tag but contained in other copies of** database

Patent Assignee: PALMSOURCE INC (PALM-N)
Inventor: DUGGARAJU R; GOEPPINGER C; JARVINEN B; MCCAW K
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
US 6711578	B1	20040323	US 2001764524	A	20010117	200427	B

Priority Applications (No Type Date): US 2001764524 A 20010117
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 6711578	B1		17	G06F-017/30	

Abstract (Basic): US 6711578 B1

NOVELTY - The edited and modified records are identified by comparing records of focus copy against **records** having **same** identification tag but contained in other copies of **database** . The record indicated as deleted is **removed** and records indicated as modified are modified to all copies of **database** . The cycle is repeated until all the copies of **database** is processed as focus copy.

DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for computer system.

USE - For synchronizing multi-copies of **database** in network environment including electronic devices like personal digital assistant (PDA), palm top computer e.g. for local area network (LAN), wide area network (WAN), Internet.

ADVANTAGE - Reduces processing time, since any record that is already processed by another focus **database** is skipped and new **database** is processed as the focus **database** . Provides the synchronization of an upwardly increasable number of copies of multiple **databases** without an attendant exponential increase in the amount of time and resources required.

DESCRIPTION OF DRAWING(S) - The figure shows the flowchart explaining method of synchronizing multi-copies of **database** .

pp; 17 DwgNo 9/10

Title Terms: MULTI; COPY; SYNCHRONISATION; METHOD; **DATABASE** ; IDENTIFY; EDIT; MODIFIED; RECORD; COMPARE; RECORD; FOCUS; COPY; RECORD; IDENTIFY; TAG; CONTAIN; COPY; **DATABASE**
Derwent Class: T01
International Patent Class (Main): **G06F-017/30**
File Segment: EPI

18/5/3 **(Item 3 from file: 350)**
DIALOG(R)File 350:Derwent WPIX

015250620 **Image available**
WPI Acc No: 2003-311546/200330
XRPX Acc No: N03-248004
Duplicate invoices **identifying method for electronic payment system, involves** replacing identical **index numbered** invoices **in single** invoice **and eliminating** compared invoices **including** replaced **invoices from** database
Patent Assignee: INT BUSINESS MACHINES CORP (IBMC)
Inventor: CALKINS W D; DONNELLY R A; MURPHY J M; VANLONE J W
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
US 20020194174	A1	20021219	US 2001832572	A	20010411	200330 B

Priority Applications (No Type Date): US 2001832572 A 20010411
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 20020194174	A1		15	G06F-007/00	

Abstract (Basic): US 20020194174 A1
NOVELTY - The **invoices** with **same** index numbers, loaded in the **database** at different timings are **replaced** in a single invoice. The report corresponding to comparison result of invoices, including **replaced** invoices is generated based on which the invoices judges to have compared are eliminated from the **database** .
DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following:
(1) Evaluated invoiced documents report providing method;
(2) Possible **duplicate invoices** packets capturing method;
(3) Program storage device storing instructions for identifying **duplicate records** ;
(4) Possible **duplicate invoices** capturing system; and
(5) Computer program element for capturing packets of possible **duplicate invoices** .
USE - For identifying **duplicate invoices** for payment, credit, goods services, among several systems such as electronic payment or enterprise resource planning (ERP) systems, multiple accounts payable (A/P) systems.
ADVANTAGE - From the invoices left after eliminating the compared **databases** , it is easy to identify and take action on **duplicate invoices** prior to payment.
DESCRIPTION OF DRAWING(S) - The figure shows a system diagram of the **duplicate invoices** identifying system.
pp; 15 DwgNo 1/5
Title Terms: DUPLICATE; INVOICING; IDENTIFY; METHOD; ELECTRONIC; PAY; SYSTEM; **REPLACE** ; IDENTICAL; INDEX; NUMBER; INVOICING; SINGLE; INVOICING ; ELIMINATE; COMPARE; INVOICING; **REPLACE** ; INVOICING; **DATABASE**
Derwent Class: T01; T05
International Patent Class (Main): **G06F-007/00**
File Segment: EPI

18/5/4 (Item 4 from file: 350)
DIALOG(R)File 350:Derwent WPIX

015227576 **Image available**
WPI Acc No: 2003-288489/200328
XRPX Acc No: N03-229320
Matching information merge and data trees prune method for world wide web, involves applying merge document **to identified** matching documents **within source documents**
Patent Assignee: INT BUSINESS MACHINES CORP (IBMC)
Inventor: MYLLYMAKI J P
Number of Countries: 001 Number of Patents: 002
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
US 20020188598	A1	20021212	US 2001834965	A	20010412	200328	B
US 6757678	B2	20040629	US 2001834965	A	20010412	200443	

Priority Applications (No Type Date): US 2001834965 A 20010412
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 20020188598	A1		15	G06F-017/30	
US 6757678	B2			G06F-017/30	

Abstract (Basic): US 20020188598 A1
NOVELTY - Two or more source documents that share a similar data structure are identified. **Matching documents** that relate to the same configurable entity within the two or more source documents are identified. A merge document is applied to the **matching documents** to merge the **matching documents** into a resultant document, and to prune the data tree of the resultant document.
DETAILED DESCRIPTION - An INDEPENDENT CLAIM is included for software program product for merging matching information and pruning data trees.
USE - For world wide web (WWW).
ADVANTAGE - Provides a way of retrieving sets of individual web pages from web sites and locally merging the data. Enables the user to obtain logical tree data structure where redundancies have been **removed** . Enables the user to bypass the built-in restrictions in product **databases** to effectively mine the data for information. Permits comparative analysis of the data.
DESCRIPTION OF DRAWING(S) - The figure shows a schematic view of an operation environment utilizing data trees automated merging and pruning system.
pp; 15 DwgNo 1/8
Title Terms: MATCH; INFORMATION; MERGE; DATA; TREE; PRUNE; METHOD; WORLD; WIDE; WEB; APPLY; MERGE; DOCUMENT; IDENTIFY; MATCH; DOCUMENT; SOURCE; DOCUMENT
Derwent Class: T01
International Patent Class (Main): **G06F-017/30**
International Patent Class (Additional): **G06F-017/24**
File Segment: EPI

18/5/12 **(Item 12 from file: 350)**
DIALOG(R)File 350:Derwent WPIX

011565014 **Image available**
WPI Acc No: 1997-541495/199750
Related WPI Acc No: 2003-261399
XRPX Acc No: N97-450829

Census registration report force entry method for birth report, death report, marriage registration - involves using terminal equipment to replace **data, acquired from either census registration** database **or resident** recording database **and** matched **with data item input to search key, into suitable data entry**

Patent Assignee: HITACHI LTD (HITA)
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 9259183	A	19971003	JP 9669503	A	19960326	199750	B

Priority Applications (No Type Date): JP 9669503 A 19960326
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 9259183	A		7	G06F-017/60	

Abstract (Basic): JP 9259183 A

The method involves acquiring data which matches with data item input to a search key through a terminal equipment (13), by searching a census-registration **database** (12) based on the search key.

When the data corresponding to the search key are not found in the census-registration **database** , the data are acquired by searching a resident recording **database** (15) which stores resident information. The terminal equipment then **replaces** the acquired data in suitable data entry.

ADVANTAGE - Reduces load of data item input even when there are no data applicable to census-registration **database** since resident recording **database** can be searched.

Dwg.1/6

Title Terms: REGISTER; REPORT; FORCE; ENTER; METHOD; BIRTH; REPORT; DEAD; REPORT; REGISTER; TERMINAL; EQUIPMENT; **REPLACE** ; DATA; ACQUIRE; REGISTER ; **DATABASE** ; RESIDENCE; RECORD; **DATABASE** ; MATCH; DATA; ITEM; INPUT; SEARCH; KEY; SUIT; DATA; ENTER
Derwent Class: T01
International Patent Class (Main): **G06F-017/60**
International Patent Class (Additional): **G06F-017/30**
File Segment: EPI

18/5/16 (Item 16 from file: 350)
DIALOG(R)File 350:Derwent WPIX

010632613 **Image available**
WPI Acc No: 1996-129566/199613
XRPX Acc No: N96-108864

Graphic data updating system for e.g. utility company map data - compares update records of graphic data held at sub-station and corresponding data held at base station and replaces older data

Patent Assignee: TOKYO GAS CO LTD (TOLG)
Inventor: ISHIKAWA Y; SATO H; YONEYAMA K

Number of Countries: 018 Number of Patents: 009
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
WO 9604606	A1	19960215	WO 95JP1493	A	19950727	199613	B
JP 8044768	A	19960216	JP 94197531	A	19940729	199617	
JP 8076685	A	19960322	JP 94230450	A	19940831	199622	
JP 8077202	A	19960322	JP 94230451	A	19940831	199622	
EP 722144	A1	19960717	EP 95926508	A	19950727	199633	
			WO 95JP1493	A	19950727		
US 5794258	A	19980811	WO 95JP1493	A	19950727	199839	
			US 96615245	A	19960506		
EP 722144	A4	19971015	EP 95926508	A	19950727	199840	
EP 722144	B1	20011031	EP 95926508	A	19950727	200169	
			WO 95JP1493	A	19950727		
DE 69523553	E	20011206	DE 623553	A	19950727	200203	
			EP 95926508	A	19950727		
			WO 95JP1493	A	19950727		

Priority Applications (No Type Date): JP 94230451 A 19940831; JP 94197531 A 19940729; JP 94230450 A 19940831
Cited Patents: 03Jnl.Ref; JP 5108729; JP 5233770; JP 6019969; JP 6325139; EP 511010
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
WO 9604606	A1	J	63	G06F-017/30	
Designated States (National): US					
Designated States (Regional): AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE					
JP 8044768	A		12	G06F-017/30	
JP 8076685	A		14	G09B-029/00	
JP 8077202	A		13	G06F-017/30	
EP 722144	A1	E	43	G06F-017/30	Based on patent WO 9604606
Designated States (Regional): DE FR GB					
US 5794258	A			G06T-001/00	Based on patent WO 9604606
EP 722144	A4			G06F-017/30	
EP 722144	B1	E		G06F-017/30	Based on patent WO 9604606
Designated States (Regional): DE FR GB					
DE 69523553	E			G06F-017/30	Based on patent EP 722144 Based on patent WO 9604606

Abstract (Basic): WO 9604606 A

Graphic data is read from optical disc (19) at a sub-station (3-1), and after designation of a search area the data is sent to the base station (1) by eg. modem (29, 33). At the base station, graphic data including the designated search area is extracted by the host computer (5) from a **database** (9).

The data and time of updating of both the extracted data (10) and the data received from the sub-station is then checked. When the time of the last update of the data extracted at the base station is more recent than that of the received data, the newer extracted graphic data is transmitted to the sub-station.

USE - For providing sub-stations, portable computers etc. with the newest available data, in a system where centrally stored data is subject to regular updating.

Dwg.1/48

Title Terms: GRAPHIC; DATA; UPDATE; SYSTEM; UTILISE; COMPANY; MAP; DATA; COMPARE; UPDATE; RECORD; GRAPHIC; DATA; HELD; SUB; STATION; CORRESPOND; DATA; HELD; BASE; STATION; **REPLACE** ; DATA

Derwent Class: P85; T01

International Patent Class (Main): **G06F-017/30** ; G06T-001/00; G09B-029/00

International Patent Class (Additional): G01C-021/20; **G06F-013/00** ; **G06F-015/00** ; **G06F-017/50** ; G08G-001/09; G08G-001/0969

File Segment: EPI; EngPI

18/5/24 **(Item 24 from file: 347)**
DIALOG(R)File 347:JAPIO

07321269 **Image available**
METHOD FOR GENERATING LIST OF RETRIEVAL RESULT OF RELATIONAL **DATABASE** , RETRIEVAL DEVICE FOR RELATIONAL **DATABASE** , AND RECORDING MEDIUM

PUB. NO.: 2002-189756 [JP 2002189756 A]
PUBLISHED: July 05, 2002 (20020705)
INVENTOR(s): MAEDA KOJI
FUJIYAMA EIKO
APPLICANT(s): FUJITSU LTD
APPL. NO.: 2000-389112 [JP 2000389112]
FILED: December 21, 2000 (20001221)
INTL CLASS: **G06F-017/30**

ABSTRACT

PROBLEM TO BE SOLVED: To provide a system which is easy to read through without any awareness of hierarchy as a system which manages a series of volumes of materials (book, etc.), through a relational **database (RDB)**.

SOLUTION: An identification key item is specified among a plurality of items (field) included in the **RDB** and when retrieval from the **RDB** is performed, **records** having **matching** data by items as to the previously specified identification key item are detected from extracted records to extract a plurality of records as ones in the same series and generate series information, which is entered into a list of retrieval results while generated by having records **replaced** with a plurality of records included in the series.

18/5/30 **(Item 30 from file: 347)**
DIALOG(R)File 347:JAPIO

05704835 **Image available**
DATABASE UPDATING METHOD

PUB. NO.: 09-319635 [JP 9319635 A]
PUBLISHED: December 12, 1997 (19971212)
INVENTOR(s): WATANABE SHINICHI
APPLICANT(s): MATSUSHITA GRAPHIC COMMUN SYST INC [330729] (A Japanese Company or Corporation), JP (Japan)
APPL. NO.: 08-136233 [JP 96136233]
FILED: May 30, 1996 (19960530)
INTL CLASS: [6] **G06F-012/00 ; G06F-017/30**
JAPIO CLASS: 45.2 (INFORMATION PROCESSING -- Memory Units); 45.4 (INFORMATION PROCESSING -- Computer Applications)

ABSTRACT

PROBLEM TO BE SOLVED: To accelerate batch processing or update processing itself of records in a **database** by performing margin processing to the received update **records** having the **same** ID, and updating the records registered in the **database** while using the records after the margin processing.

SOLUTION: A batch file managing part 20 is activated and it is checked whether the **record** of the **same** ID as a record becoming the object of update processing is defined as the object of new registration or not (exists in an update file 26 or not). When the **record** of the **same** ID exists in the update batch file 26, the batch file managing part 20 performs the marging processing for the unit of a field to the update record and the update record existent in the update batch file 26 through an update batch file access part 25, generates the new update record, deletes the existent update record and **replaces** the generated new update record with the deleted update record.

```
Set     Items   Description
S1     163084   DATABASE? OR DATABANK? OR DB OR DBMS OR RDB OR OODB OR DBS
            OR DATA()(BASE? OR BANK?)
S2    1144561   RECORD? OR DOCUMENT? OR INVOICE? OR BILL OR BILLS OR RECEI-
            PT?
S3    1648187   DUPLICATE OR IDENTICAL OR DUPLICATES OR SAME OR EQUIVALENT
S4    2716345   REMOVE? OR REPLAC? OR SUBSTITUT? OR SWITCH? OR SWAP?
S5    2724453   DATE OR DATES OR DATESTAMP?  OR INDICIA? OR TIMESTAMP? OR -
            TIME OR CALENDAR?
S6     239544   BLANK? OR NULL OR EMPTY OR VOID?
S7    1302872   MATCH? OR COMPARE? OR EQUAL OR EQUIVALENT?
S8      15609   S2(2N)(S3 OR S7)
S9        992   S1 AND S8
S10        64   S4 AND S9
S11       296   S1 AND S5 AND S6
S12     12638   S5(2N)(INSERT? OR FILL OR FEED OR POPULAT?)
S13         1   S11 AND S12
S14        80   S12(5N)S6
S15        46   (S10 OR S13 OR S14) AND IC=G06F?
S16        36   S15 NOT AD>20010412
S17        36   IDPAT (sorted in duplicate/non-duplicate order)
S18        35   IDPAT (primary/non-duplicate records only)
S19    630781   BLEND OR BLENDS OR BLENDING OR MERGE? OR MERGING OR COMBINE
              OR COMBINES OR COMBINING OR APPEND? OR MIX OR UPDAT?  OR UP()-
             (DATE OR DATING OR DATES)
S20        14   S10 AND S19
S21         1   S20 NOT S15
S22       756   S8 AND S19
S23       147   S1 AND S22
S24       131   S23 AND IC=G06F?
S25       118   S24 NOT (S10 OR S13 OR S14)
S26         9   S25 AND IC=G06F-007?
File 347:JAPIO Nov 1976-2004/May(Updated 040903)
         (c) 2004 JPO & JAPIO
File 350:Derwent WPIX 1963-2004/UD,UM &UP=200459
         (c) 2004  Thomson Derwent
```

26/5/3 **(Item 3 from file: 350)**

DIALOG(R)File 350:Derwent WPIX

015608507 **Image available**

WPI Acc No: 2003-670664/200363

XRPX Acc No: N03-535495

Unique object record identification using rule analyzer system for healthcare organization, involves determining efficiency of exact match and probabilistic search rules, to accordingly adjust rules in descending order

Patent Assignee: ECLIPSYS CORP (ECLI-N)

Inventor: TIFFT W W

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
US 20030120652	A1	20030626	US 99160717	P	19991019	200363 B
			US 2000692433	A	20001019	
			US 2003349304	A	20030121	

Priority Applications (No Type Date): US 99160717 P 19991019; US 2000692433 A 20001019; US 2003349304 A 20030121

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 20030120652	A1		19	G06F-007/00	Provisional application US 99160717
					Div ex application US 2000692433

Abstract (Basic): US 20030120652 A1

NOVELTY - The user defined probabilistic search rules are executed to search a unique object record in a **database** , if exact match search rules do not retrieve **identical** object **records** . The user selected object record is **updated** with new attributes in real-time. The efficiency of exact match and probabilistic search rules are determined, to accordingly adjust the rules in descending order.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) unique object record identifying system;

(2) rules analysis method; and

(3) rules analyzer system.

USE - For identifying an object record, using a rules analyzer system (claimed) in healthcare organization.

ADVANTAGE - Efficiently evaluates the efficiency and reordering of exact match and probabilistic search rules, thus maintaining a set or rules to locate the desired record in an efficient manner.

DESCRIPTION OF DRAWING(S) - The figure shows the display screen of a rule generator.

pp; 19 DwgNo 2/9

Title Terms: UNIQUE; OBJECT; RECORD; IDENTIFY; RULE; ANALYSE; SYSTEM; ORGANISE; DETERMINE; EFFICIENCY; EXACT; MATCH; PROBABILITY; SEARCH; RULE; ACCORD; ADJUST; RULE; DESCEND; ORDER

Derwent Class: T01

International Patent Class (Main): **G06F-007/00**

File Segment: EPI

26/5/8 **(Item 8 from file: 350)**
DIALOG(R)File 350:Derwent WPIX

010800295 **Image available**
WPI Acc No: 1996-297248/199630
XRPX Acc No: N96-250116

Data sorting system - has merge **part which sequentially outputs sorted record to work-file without allowing record to participate again in sorting process**

Patent Assignee: NEC CORP (NIDE)
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 8129478	A	19960521	JP 94290451	A	19941031	199630	B

Priority Applications (No Type Date): JP 94290451 A 19941031
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 8129478	A		21	G06F-007/24	

Abstract (Basic): JP 8129478 A

The sorting system has a management table (10) which sets an equivalent key flag for each record of the **database** . At the time of sorting, the records are read into a memory (1) sequentially. When a record is sorted out, a part (2) sets its equivalent key flag.

A **merge** part (4) obtains the **record** whose **equivalent** key flag is set and is output sequentially into a work file (7). Along with the record in the work-file, its equivalent key flag information is added by a pre sorting part. The sorting process is continued for the rest of the records in the string.

ADVANTAGE - Reduces number of comparison process for sorting. Reduces sorting time.

Dwg.1/11

Title Terms: DATA; SORT; SYSTEM; **MERGE** ; PART; SEQUENCE; OUTPUT; SORT; RECORD; WORK; FILE; ALLOW; RECORD; PARTICIPATING; SORT; PROCESS
Derwent Class: T01
International Patent Class (Main): **G06F-007/24**
File Segment: EPI

26/5/9 **(Item 9 from file: 350)**
DIALOG(R)File 350:Derwent WPIX

010654050 **Image available**
WPI Acc No: 1996-151004/199615
Related WPI Acc No: 1997-052579; 1998-145142
XRPX Acc No: N96-126963

Merging databases **of records in parallel and redundancy checking of mailing list - comparing sorted records based on first key to each other, identifying** duplicate records **if less than number of records in** database **, storing identity, repeating sorting for second key and subjecting union of keys to transitive closure**

Patent Assignee: STOLFO S J (STOL-I)
Inventor: HERNANDEZ M A; STOLFO S J
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
US 5497486	A	19960305	US 94213795	A	19940315	199615	B

Priority Applications (No Type Date): US 94213795 A 19940315
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 5497486	A		22	G06F-007/06	

Abstract (Basic): US 5497486 A

The method involves computing a first and second key respectively for each record in each **database** by extracting a portion of a first and second field. The records in each **database** are parallel **merge** sorted using the first and second keys respectively.

The method also involves comparing to each other, on a first and second group of processors, a predetermined number of sequential records sorted according to the first and second key respectively to determine if one or more of the **records match** . The identifiers for any **matching records** of both keys are stored.

A union of the stored identifiers is created and subjected to transitive closure, where, in each group of processors for each **database** , N is a number of records, P is a number of processors. Each processor p, 1 at most p at most P, is able to store M+w records, where w is a size of a **merge** phase window and M is a blocking factor.

In each group of processor for each **database** P is less than N, MP is less than N and ri represents record i in a cluster, 0 at most i at most MP-1. Each comparing step involves dividing the sorted **database** into N/MP clusters, processing each of the N/MP clusters in turn by providing each processor p with records r(p-1)M, . . . , rpM-1, . . . , rpM+w-2, for 1 at most p at most P. The **matching records** are searched independently at each processor using a window of the size w. Finally, the processing step is repeated for a next cluster of records.

ADVANTAGE - Data clustering reduces complexity to linear time making multiple runs followed by transitive closure feasible and efficient. Large **databases** are accommodated which uses parallel and distributed computing to achieve efficient performance with acceptable cost.

Dwg.7/8

Title Terms: **MERGE** ; RECORD; PARALLEL; REDUNDANT; CHECK; MAIL; LIST; COMPARE; SORT; RECORD; BASED; FIRST; KEY; IDENTIFY; DUPLICATE; RECORD; LESS; NUMBER; RECORD; **DATABASE** ; STORAGE; IDENTIFY; REPEAT; SORT; SECOND ; KEY; SUBJECT; UNION; KEY; CLOSURE
Derwent Class: T01
International Patent Class (Main): **G06F-007/06**
International Patent Class (Additional): **G06F-007/14** ; **G06F-007/20**
File Segment: EPI